【知识福利社】你的私人书院，微信：312005650 全网热门课程众筹共享！

09｜视野：不同演化下的AI分支有哪些？

卓克·怎样用AI高效学习
今天

进入课程 >

时长：12分33秒 大小：11.48MB

00:00 / 12:33

卓克亲述

🇨🇳，你好。

欢迎回到《怎样用AI高效学习》，我是卓克。

课程的最后一讲，我想把视野扩得更大，跳出ChatGPT的束缚，为你介绍不同演化分支下的其他AI。

这里的AI，指的是利用大语言模型做出来的产品。有人认为，我要介绍New Bing、文心一言的功能、对比效果。其实不是，而且这些内容在网上本来就已经很多了，不需要我来做了。

在我看来，真正能使用大语言模型和OpenAI竞争的是那些从利益、理念上就不一样的公司。

今天，我们就来介绍一下这些分支，它们在今后的演化才是AI时代的全景图。

Anthropic公司

第一个要介绍的是Anthropic公司。它推出的产品叫Claude，也是聊天机器人的形态。不过并没有公开测试，只有很小规模的人拿到了测试资格。

Anthropic公司的创办者达里奥·阿莫迪（Dario Amodei）是物理学出身，从加州理工、斯坦福，再到普林斯顿拿到博士，全都是物理学。

他是2016年7月加入OpenAI团队的，属于元老级的，做到研发副总裁的职位后离开。GPT-2和GPT-3就是在他的带领下诞生的产品。但是，当需要微软额外投入10亿美元才能开始下一轮模型训练时，阿莫迪就已经不太同意这个发展方向了。

大家在之前已经发现了规律，就是超大规模的参数+超大规模的数据+超大规模的训练量叠加在一起，就会出现能力的“涌现”，涌现出来的能力是训练前根本无法预知的。这很好，很令人惊讶，但也很烧设备，很烧电费。阿莫迪不认为这样烧钱的方式可以持续。

而且，即便烧钱训练，也还是会不可避免的出现严重问题。

比如，AI领域有一个专业说法叫作“目标错位”（Misaligned goals）或者“奖励劫持”（Reward Haching），说的就是，经过超大规模训练后的机器人，在比较小规模的测试中能够表现得很得体；而一旦放在更大规模的真实环境里，就会逐渐偏离任务目标。

举个例子：

微软当年公布聊天机器人Tray时，最初设定的目标是“参与度最大化”。在微软的模型里，“参与度”当然是一个严格的数学概念，但我们大致可以理解为，尽量让对话内容吸引人继续下去。这是一个比较美好和正向的目标。

但实际上，没过几天，在成千上万人和Tray聊天以后，Tray就开始靠喷种族主义言论和飙脏话来达成目标了。这就叫奖励劫持。

说回Anthropic公司的产品Claude。拿到测试资格的人后来纷纷发布了自己的使用体验，和ChatGPT3.5做对比，PK下来的总体感觉是：

与ChatGPT相比，Claude能更明确的拒绝不恰当的请求。遇到能力范围之外的问题，Claude会主动承认自己回答不了。Claude给出的答案更长、更啰嗦，但句子衔接更自然。但在程序代码方面，Claude的性能敌不过ChatGPT，代码里有更多错误。数学、逻辑能力，二者相似。

这时候，有些人就要问了：既然是理念不同出走的，做出来的产品好像大同小异啊？

不。以聊天机器人胡说八道这个问题来说，两者的处理方式就是截然不同的。

当ChatGP出现大规模的频繁的编造虚假内容的情况的时候，工程师们只能靠再次训练才能优化这个问题。而参数规模达到几千亿后，重新训练的成本是非常高昂的。

Claude虽然也会自己瞎编一些内容，但是要纠正Claude胡说八道的问题时，并不需要重新训练，而是有很大可能直接在模型参数设置上找原因，然后针对性的调整那些参数的权重，问题就有可能得到解决。

ChatGPT为什么不能直接调整模型参数呢？

首先，它的参数规模1750亿，太大了。

其次，深度神经网络具有一个不可避免的性质——不可解释性。你不知道输入了成吨的素材训练之后，它对某几个问题的输出结果为什么会如此奇葩，你无法解释。你只能说，当下这1750亿个参数值分别等于多少多少以后，把一个输入灌入这个模型后，必然会产生这样奇葩的输出。解释完毕。

你说，这也能算解释？是的，跟没解释一样。相当于问，人为什么会直立行走，解释是，人类物种和与之相关的全部环境因素综合演化的完整历程决定了人会直立行走。你看，跟没回答是一样的！

Claude使用的模型可以解释一部分结果为什么是那个样子，于是调整优化的时候就可以节省大量训练成本。

商业公司竞争，主要靠三方面——产品素质、经营成本、人员管理。在初步测试中，Claude和GPT3.5的性能表现比较接近，后续如果进步速度也类似的话，Claude在成本上的优势就太大了。

因为它具有一定的可解释性，所以可以预设更多的条条框框给模型，让它具有人们指定的价值观和偏好。而这一点，既可以用来预防虚假信息，也可以用来配合某些政府控制大语言模型的海量输出。

顺便说一句，马斯克的TruthGPT也是选择往这个方向发展的，尽管他也是OpenAI早期的投资人。

CovariantAI公司

第二个要介绍的是CovariantAI公司。它并不是在生成式AI领域和OpenAI竞争，却可能是在OpenAI的技术框架下对现实世界改变最大的，因为它在设计机器人之间的协作和交互。

我们知道，现在GPT-4和类似的竞争者都在往多模态输入的方向发展，也就是说，让它不但读得懂文字，也能看懂图片，甚至今后还有可能看懂视频。当这些多模态的应用成熟、移植到有行动能力的机械手臂、运货平台、机床上的时候，很多从前需要复杂编程的过程就会大大省略。

而CovariantAI要做的，就是把机械手臂、扫地机器人、运货平台、机床、人形机器人等不同形态的机器人之间的接口打通，让它们可以协作起来。

听着是不是有点科幻？离取代人力都不远了？

至于什么时候能做到，能不能做到，政府允许不允许科技往这个方向发展，都是未知数，但这个AI的演化分支是客观存在的。

现在，CovariantAI的产品都是给物流仓库使用的。大型仓库内的分拣和运输，由它们解决不同类型机器人之间的互动。

今后可能会出现一种AI机器人的通用数据接口标准，类似于IT时代的USB口那样，功能和设计差异很大的工具都可以通过USB口实现数据和指令的传输。今后十万八千种AI机器人，很可能也有这样的接口标准。

其实，这条发展方向本来就是OpenAI公司自己从前的规划，只不过在2021年后，它把重点放在了获取数据和训练大语言模型上，机器人方面的项目就停掉了。

而CovariantAI公司的CEO、CTO，曾经全都是OpenAI的技术核心，他们把这个发展方向延续了下去。

ADEPT公司

第三个要介绍的是ADEPT公司。它也不是要在算法和模型设计上与OpenAI竞争，而是要把API接口最大化。

ChatGPT结合API接口这件事，放在2023年3月份已经不稀奇了。因为OpenAI自己就主动公布了13个插件，它们都可以利用GPT-4的API接口。但ADEPT是从2022年1月就开始着手在做了。

它的终极目标是，建立所有软件与人工智能平台的接口，比如图片处理的Photoshop、视频编辑的PR、音频处理的Audition等，至于接入哪个人工智能平台，也许是OpenAI的，也许是谷歌的，也许是今后的Anthropic的。

而今天我们见到的开放出来的API接口都是各家分别推出的。

比如，你要用AutoGPT，起码得把谷歌搜索的API、Pinecone的API、GPT-3.5或者4的API拿到，这样才能把AutoGPT部署到本地。

或者哪天Adobe全家桶、微软全家桶和ChatGPT打通了，也依然需要我们单独购买、下载AI版本的全家桶软件。

如果ADEPT成熟了，苹果和谷歌的软件商店，还有各种作为流量入口的工具，比如浏览器的访问，那个必要性，对用户来说都会大幅下降。因为那时候，ADEPT的通用型人工智能界面会把零零散散的软件和API接口集成在一起，我们直接和ADEPT的软件界面打交道就行了。

总结

以上就是我们现在可以看到的大语言模型下今后可能有大发展的不同分支。

你发现没有，绝大部分是OpenAI早期的元老出走后创办的。直接原因其实和2019年底，OpenAI为了找足够的钱训练GPT-3，不得不把当时还是非营利性组织的OpenAI改成商业性质的公司有关。而这次大变动，也动摇了原来的利益分配格局。

虽然那些公开宣布离开OpenAI的人总是强调是自己的研发理念不同，但根源上说，还是因为利益冲突。

最后，我们说说这一波大语言模型的影响范围。

有人对此非常担忧，比如AI会接管人类，AI开始自己优化自己，或者世界上最终只有几百个尖端的人处于AI之上，剩下的人都是《人类简史》里描述的那种无用阶级。

这些说法多少都科幻化了。实际上，很多人没有意识到，也很难意识到的一个重要结论是——人类的大部分知识是非语言的。支撑我们正常生活的很多东西与语言无关。

有很多更基础的能力，并不是语言，比如自发形成的分类能力、移动物体的追踪能力、识别物体之间互相遮挡的关系的能力。

我们都知道，几个月的婴儿就可以分出自行车和汽车，并且追踪它们的位置，而那个时候，他们并不具备表达和认知自行车和汽车的语言能力。

现在，AI科学家还不知道怎么给纯粹的非语言能力做训练。并且，以文本生成方式发展出来的续写能力，也并不能很好的迁移到其他数据类型上。这一方面因为数据量过大，另一方面也因为，关于神经网络的工作方式，是不是还需要添加很多额外的工作机制才能让模型输出更好的结果，一切都在未知中。

但业内人士大都认为，要展现更全面的近乎于人的能力，当前的模型不太可能胜任。

即便如此，这波技术变革也依然是异乎寻常的。

一年后，当你再次回看这个课程，肯定会发现，所有软件的使用、API接口的申请全都过时了，功能更全、更强大、使用更简单的产品早已出现。

但起码当下，通过这门课，希望你提前自己的竞争对手熟悉AI，用上AI。先发优势是很重要的。后续有这方面的动态，我也会在知识城邦第一时间跟你报告。

好，我们的课程就到这里。《科技参考3》，不见不散。

划重点

1. 真正能使用大语言模型和OpenAI竞争的是从利益、理念上就不一样的公司。

2. 人类的大部分知识是非语言的。支撑我们正常生活的很多东西与语言无关。

3. 要展现更全面的近乎于人的能力，当前的模型不太可能胜任。

内容禁止商业，仅供社群内部免费交流使用
【知识福利社】微信:312005650

首次发布：今天

上一篇 08｜自动执行：下一代的GPT长什么样？

赞赏

1 人已赞赏

可免费赠送给20位朋友

微信 朋友圈

用户留言

最新 最热 只看作者回复

我是运营 3小时前 + 关注

卓老板为什么chatGPT爆火以后，国内各种科技公司开发为什么能在那么短时间开发出或者正在开发马上就能上市自己AI？没技术门槛吗？最大的成本是训练成本吗？如果是这个训练成本这其中电费，硬件支出是最高的吗？那硬件没要求吗？咱计算机领域不是被掐脖子了吗？影响不到吗？大量优质的有效训练素材是成本吗？按照我粗糙理解，我只要是个半吊子程序员，用chatGPT，在开发一个AI，然后用自己服务器训练他，只要求能满足我个人需求，我是不是也能开发一个AI？我这个问题是不是算上段问题的答案？

收起

转发 2 赞 分享

作者回复：
差不多，实在不行先把名字起出来，悟道，参透之类的，先训练出一个GPT-1.5水准，再开发布会

郑跃葵 4小时前 + 关注

感谢卓老板的先一步探索，我们后续课程不见不散

转发 评论 赞 分享

囚之 4小时前 + 关注

卓老板，这些科技公司不公开测试是因为有什么顾虑吗？现在ChatGPT大火，按照描述Claude不会逊色于ChatGPT，却鲜少有人知道。ChatGPT如此火爆，它的对手们不怕失去了先发优势吗？

转发 3 1 分享

作者回复：
可能成本太高了

緋剑心ㄱ 5小时前 + 关注

催更科技参考3，放出预购也行啊

转发 评论 2 分享

一碗水饺 5小时前 + 关注

老板的节目回来了，我又回归零点守候节目的状态。
快速扫完了文稿，看到最后，脑中响彻孙楠的不见不散😊

1 1 赞 分享

作者回复：
不，《科技参考3》到时会改时间，不会每天0点，咱们还改成19:17，老听众都知道这个梗吧。

以上留言由作者筛选显示

【知识福利社】你的私人书院，微信：312005650 全网热门课程众筹共享！